Microstructure Based Indicia

Joshua R. Smith
Escher Laboratories
101 Main Street
Cambridge, MA 02142
jrs@escher-labs.com

Andrew V. Sutherland
Escher Group, LTD
101 Main Street
Cambridge, MA 02142

Abstract

Value indicia are marks printed on paper that prove that a particular service has been paid for, one prevalent example being the postal meter mark. This paper explores systems for producing and verifying value indicia that prevent the indicium from being fraudulently copied, by binding the indicium to unique and typically random characteristics of the physical substrate on which it is printed. Thus a copy of a valid indicium will not itself be valid, because of the mismatch between the substrate microstructure on the original and the copy. In addition to the direct benefits associated with reducing fraud due to copying of valid indicia, such a system has a number of collateral security and other benefits that will be explained.

Note: U.S. and international patents are pending on the technology described in this paper.

Background: Value Indicia

Value and authentication indicia are playing an increasingly important role in our information-driven economy. A value indicium is typically a mark printed on a piece of paper that indicates that a particular service has been paid for. One prevalent example is the "franking" or postal meter mark. If such a mark appears on a mailpiece, the postal service will agree to deliver it, because the mark provides evidence that the user has paid for its services. Other examples of value indicia include currency itself, money orders, and tickets for cultural events or transportation. Authentication indicia indicate that a particular document is valid, not a forgery or altered version of a genuine document. For example, the various stamps, seals, and other indications of authenticity on a passport or driver's license would constitute authentication indicia.

Appearance Based Indicia

Historically, the validity of an indicium was based on its *appearance.* If the mark from a postage metering machine looked genuine, it would be accepted as genuine by the postal service. The security of appearance-based indicia is often based on difficulty of printing some of the features. Improvements in copying machines, and also improved personal computer peripherals such as scanners and printers, are rendering counterfeiting of appearance based indicia easier. The increasing ease of counterfeiting traditional indicia, plus the possibility of producing indicia from ordinary PC printers, has lead to increased interest in *information-based,* rather than appearance-based, indicia.

Information Based Indicia

The validity of information-based indicia is established cryptographically: information such as the postage date, meter number, meter sequence number, source address and destination address are cryptographically signed, and this information, including the signature, constitute the indicium. If a valid indicium is copied onto a second envelope and then the address is modified in a fraudulent attempt to send an additional letter at no additional cost, the signature bits will no longer validate with the address. In this way, information-based indicia prevent some types of fraud.

However a feature that information-based indicia share with all electronic value systems, including digital cash systems, is that unless additional measures are taken, a valid indicium can be copied and then spent many times over. So in the information-based indicia scheme described above, several pieces of mail could be sent to the same destination for the price of a single

piece of mail, and unscrupulous direct marketers, for example, would have the means and the motive to commit this sort of fraud. This type of fraud is referred to as double spending.

Double Spending
There are in fact at least two distinct types of double spending fraud that may occur in a postal indicia system. The first will be termed "photocopy double-spending," which occurs when a valid indicium is copied onto another piece of paper, for example with a photocopier machine. Photocopy double spending does not actually require the use of a copying machine, but includes any fraud in which the indicium itself is copied. The second type of double spending fraud arises when copies are made of the bits or digital tokens that authorize the production of valid indicia, rather than copies of indicia themselves. This second type of double spending will be will be termed "meter tampering double spending." This term denotes tampering with the postage meter (which could be software that is not always connected to a network, for example the Riposte Dispatch Universal Postal Client product of Escher Group, LTD, described in [2]) in order to produce more than one printed indicia for each legitimately authorized one. In a system such as the Universal Postal Client, this could occur by making copies of the bits on the computer's hard drive that authorize the production of valid indicia.

To avoid the double-spending problems that are fundamental to information-based value schemes, including both electronic postage and digital cash schemes, a database for tracking spent value tokens must typically be maintained. Then if a value token is fraudulently spent a second time, a search of the database will detect this, and prevent the transaction from being authorized.

Shortcomings of Information Based Indicia
Existing information-based indicia schemes have several major shortcomings. Because of the high cost of maintaining and querying a spent value database, some systems, such as the USPS's IBIP system (Information Based Indicia Program [3]) require a large amount of information, such as the destination address, to be

represented in the indicia, along with a digital signature of that information. But storing large amounts of information on the envelope face is costly, and yet provides little protection against photocopy double spending.

This paper describes a *microstructure based indicia* system that (1) offers resistance to photocopy double-spending without requiring an online database for verification, (2) facilitates enforcement of meter tampering double spending, and (3) decreases the amount of ancillary information (such as source and destination address) that must be encoded on the envelope.

Microstructure Based Indicia
The enhanced security described above is achieved by binding the indicium to a particular physical object, often a piece of paper, by utilizing physical features of the paper that vary randomly from one piece of paper to the next, and can be reliably sensed but cannot reliably be reproduced, for example because of their small size or three-dimensional structure. The system is analogous to a biometric identification system, in which a Personal Identification Number (PIN) is checked against a unique physical characteristic of the person presenting the number. The PIN number itself is analogous to the indicium, and the paper microstructure is analogous to the physical property on which the biometric is based---the fingerprint, or iris pattern, for example.

In the example implementation reported here, a region of paper is illuminated to make visible the randomly varying microscopic texture that is an inevitable result of the paper production process. An image of this texture is then captured with a video camera and digitized. Figures 1 and 2 show texture images for two different paper regions. A variety of other physical uncopiability primitives may also be contemplated.

The square box and the two spots in the image are marks printed on the paper to identify the region of the paper that will constitute the texture characteristic. The round spots in the upper left and lower right

of the box are registration marks used by the system. When a texture is initially "enrolled" in the system, and when a candidate texture is later presented to the system for verification, the image is rotated and translated to bring the registration spots into standard locations. The paper texture is sampled in the square region defined by the registration spots.

A "texture hash string" is then generated from the sampled region. The texture hash string is a number computed from the texture image with the property that when the same texture region is presented to the system, the same (or nearly the same) "texture hash" will be generated.

Creating and Validating Indicia
To create an indicium, a piece of paper is first enrolled in the system as follows. The registration marks are initially printed, the texture region is scanned, and then the texture hash is computed and printed elsewhere on the paper in machine readable form. The machine readable texture hash string forms the first part of the indicium. The other part of the indicium is a digital signature of the texture hash string, produced using whatever secure postage metering system (either hardware or software) comprises the rest of the system. Essentially the user must pay for the privilege of having the postal authority (or its hardware or software representative, the meter) digitally sign the texture hash string.

To validate an indicium, the machine readable indicium value, including the texture hash string and the signature, are read. First the signature is checked against the hash string. This cryptographic portion of the verification process would also occur in an ordinary information based indicia scheme. Next comes the test that is unique to the microstructure based indicia system: the texture of the paper is scanned, a candidate texture hash string is computed, and then compared to the original texture hash string that comprised the first portion of the machine readable indicium. The comparison generates a correlation score. The correlation score is then thresholded in accordance with the desired false acceptance / false reject rates for the

system, just as in a biometric identification system.

Texture hash functions
A trivial example of a texture hash is the identity function: the image is its own hash. For practical purposes, it will usually be desirable to employ a hash function that compresses the texture data more significantly. A binary thresholded version of the image would suffice as a texture hash, though typically much more compression will be desired. A version of the image compressed with a lossy compression algorithm, such as JPEG or wavelet compression, would also suffice. An approach used by Daugman to derive a hash from the texture of the human eye's iris (an "IRIS code") for biometric purposes is to use binary thresholded Gabor wavelet coefficients [1]. For practical authentication implementations, a small number of coefficients from a discrete cosine transform, discrete fourier transform, or wavelet transform may suffice.

Correlation Score
The texture hash string, whether it is the full image, a compressed version of the image, or just a few features extracted from the image, can be viewed as a feature vector. After normalizing them, a pair of feature vectors can be compared by taking their inner product, which measures the angle between the vectors in the high dimensional feature space. In statistical terminology, this inner product is known as a correlation coefficient. A value near zero indicates that the vectors are uncorrelated, and thus suggests that the underlying objects (textures) are probably different. A value near +1 indicates that the textures are probably the same. As pointed out by Daugman [1], the entropy of the probability distribution describing the ensemble of texture hashes determines the security of the system. This parameter depends on the size of the texture region, on how it is hashed (particularly the length of the hash string), and on properties of the paper itself. The numerical value of this entropy can only be determined empirically by computing texture hashes for many pieces of paper.

Implementation

Our initial proof of principle implementation uses the trivial identity hash function. To enroll a piece of paper, we capture an image, automatically find the registration marks, register the image into a standard orientation based on these marks, and save this image. To verify, we grab an image of the candidate texture region, find the registration points, bring it to the standard orientation, compute the normalized cross correlation between the candidate and original image, and threshold the result to decide whether the same piece of paper has been presented.

Discussion

The use of microstructure based indicia could completely eliminate photocopy double spending, as well as photocopy counterfeiting of authentication indicia. This is desirable since photocopy fraud is so easy to commit. No technical sophistication beyond operating a photocopying or fax machine is required. Furthermore, this form of fraud does not require the cooperation of a legitimate postage meter user, or even access to a postage meter. An individual intent on committing this form of fraud can do so with complete anonymity, possibly stealing postage generated by many different meters. All that would be required is access to mailpieces bearing valid indicia and some form of copying capability.

As well as the direct benefits from the elimination of photocopy fraud, there are additional indirect benefits. If the microstructure based indicia system is used in conjunction with technology that makes meter tampering very difficult, then it would no longer be necessary to maintain a spent value database at all, because every significant form of double spending is infeasible. Verification of postage could then occur completely offline. The savings from the elimination of the networking and database infrastructure would be enormous.

In a microstructure based system that used offline software metering, a spent value database would still be needed to check for meter tampering, but the number of database queries could be reduced, and the elimination of photocopy double spending

strengthens the incentives against meter tampering fraud, as we will explain now.

When double spending is detected by a search of the spent value database, the envelope in question can also be subjected to the test for photocopy double spending. If it fails this test, then presumably photocopy double spending has occurred. On the other hand, if it passes this test, then meter tampering double spending has definitely occurred. Assuming that the postal meter number is recorded in the indicia, the owner of that meter number can then be confronted with strong evidence that they are responsible for meter tampering, assuming that the design of the system makes impersonating another meter infeasible. In the absence of the test for photocopy double spending, the owner of the meter could plausibly claim that while the envelope was out of their control (perhaps while en route to the postal service) a third party had committed photocopy double spending. Thus the reliable means of detecting photocopy double spending also provides stronger counter-incentives for those who would commit meter tampering.

Conclusion

For value indicia as inexpensive as postage, the most cost effective postage enforcement program will probably be statistical, performing various tests for evidence of postage at various rates. The use of physical uncopiability as a primitive in the design of such systems can reduce the rate of spent value database querying, reduce the size of the spent value database, or even eliminate the database and associated network infrastructure altogether. This reduces the cost of enforcement, or allows greater levels of enforcement for the same cost as ordinary information based indicia systems that do not make use of physical uncopiability.

For authentication indicia, the benefits are similar, and possibly even greater. Since, unlike postage, items such as identity documents are normally produced at centralized, secure facilities under the control of the issuing authority, "meter tampering" is not an issue, which is the best possible case for verification: no database

and networking infrastructure is required---
verification is possibly completely offline.

References
[1] Daugman, J. (1993) High confidence
visual recognition of persons by a test of
statistical independence. IEEE Trans.
Pattern Analysis and Machine Intelligence,
15(11): 1148-1161.

[2] Escher Group, LTD Riposte Dispatch
Universal Postal Client datasheet.

[3] Information Based Indicia Program Open
System Indicium Specification. (Draft) July
23 1997. U.S. Postal Service white paper.
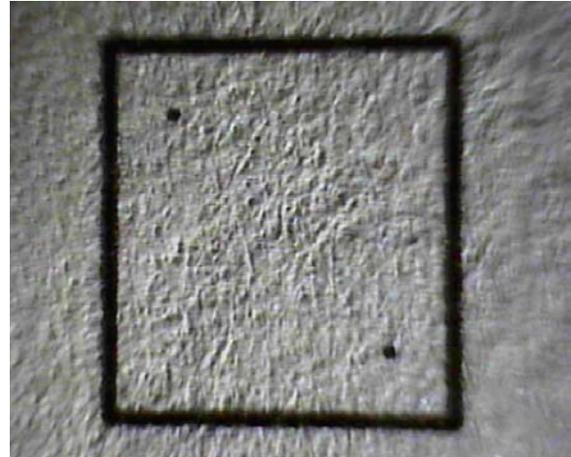http://www.usps.gov/ibip/documents/specs/i
nd72397.pdf

Figure 1. An image of paper texture. The
printed box is 8mm on a side. The round
printed spots are registration points that the
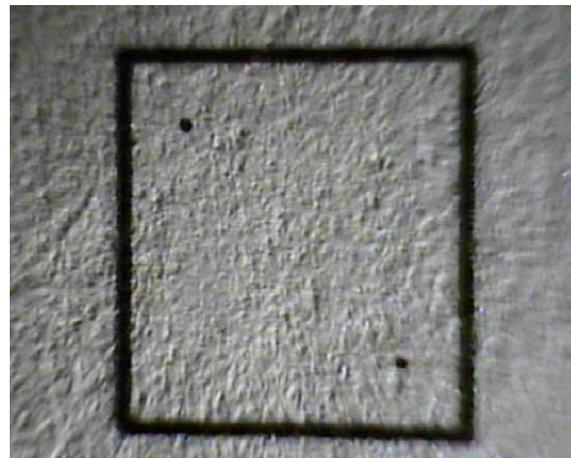system uses to bring the texture image into
a standard position and orientation.

Figure 2. A second paper texture region.